



Leveraging in big data regression

Ping Ma*

University of Georgia, Athens, Georgia, USA – pingma@uga.edu

Advances in science and technology in the past a few decades have led to big data challenges across a variety of fields. Extraction of useful information and knowledge from big data has become a daunting challenge to both the science community and entire society. To tackle this challenge requires major breakthroughs in efficient computational and statistical approaches to big data analytics.

In this talk, I will present some leveraging algorithms, which make a key contribution to resolving the grand challenge. In these algorithms, by sampling a very small representative sub-dataset using smart algorithms, one can effectively extract relevant information of vast data sets from the small sub-dataset. Such algorithms are scalable to big data. These efforts allow pervasive access to big data analytics especially for those who cannot directly use supercomputers. More importantly, these algorithms enable massive ordinary users to analyze big data using tablet computers.

Keywords: subsampling; big data; regression; least squares.